Psychological Monographs

General and Applied

No. 395
1955

Gerald F. Rabideau

Differences in Visual Acuity Measurements Obtained with Different Types of Targets

By

Gerald F. Rabideau

Purdue University

Price $1.00

Vol. 69
No. 10

Edited by Herbert S. Conrad

Published by The American Psychological Association, Inc.

Psychological Monographs:
General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

Editor

HERBERT S. CONRAD

Department of Health, Education, and Welfare
Office of Education
Washington 25, D.C.

Managing Editor

LORRAINE BOUTHILET

Consulting Editors

DONALD E. BAIER	HAROLD E. JONES
FRANK A. BEACH	DONALD W. MACKINNON
ROBERT G. BERNREUTER	LORRIN A. RIGGS
WILLIAM A. BROWNELL	CARL R. ROGERS
HAROLD E. BURTT	SAUL ROSENZWEIG
JERRY W. CARTER, JR.	ROSS STAGNER
CLYDE H. COOMBS	PERCIVAL M. SYMONDS
JOHN F. DASHIELL	JOSEPH TIFFIN
EUGENIA HANFMANN	LEDYARD R TUCKER
EDNA HEIDBREDER	JOSEPH ZUBIN

MANUSCRIPTS should be sent to the Editor.

Because of lack of space, the *Psychological Monographs* can print only the original or advanced contribution of the author. *Background and bibliographic materials must, in general, be totally excluded* or kept to an irreducible minimum. Statistical tables should be used to present only the most important of the statistical data or evidence.

The first page of the manuscript should contain the title of the paper, the author's name, and his institutional connection (or his city of residence). Acknowledgments should be kept brief, and appear as a *footnote* on the first page. No table of contents need be included. For other directions or suggestions on the preparation of manuscripts, see: CONRAD, H. S. Preparation of manuscripts for publication as monographs. *J. Psychol.*, 1948, 26, 447-459.

CORRESPONDENCE CONCERNING BUSINESS MATTERS (such as author's fees, subscriptions and sales, change of address, etc.) should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C. Address changes must arrive by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

COPYRIGHT, 1955, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Vol. 69, No. 10 Whole No. 395, 1955

Psychological Monographs: General and Applied

Differences in Visual Acuity Measurements Obtained with Different Types of Targets[1]

Gerald F. Rabideau
Purdue University

Introduction

Purpose of Study

The general aim of the present investigation was a comparison of visual acuity thresholds obtained by the same measurement method applied to several commonly used visual acuity targets. Secondary purposes were the estimation of the reliability of the thresholds and the determination of their intercorrelations. Visual acuity is herein defined as the ability of the visual sense organ to resolve images of extremely small visual stimuli.

Previous Visual Acuity Research

Since Dr. Snellen (22) first published, in 1862, his visual acuity test characters, and established one- and five-minute dimensions as a primary rule in the construction of target letters, there have been a number of attempts, both systematic and haphazard, to develop the ideal test symbol. Psychological literature abounds in reports of research on problems of visual acuity. Few, however, have systematically dealt with the evaluation and/or comparison of either new or widely used visual acuity test symbols, or targets. Recent nonoverlapping bibliographies of published papers on general visual acuity problems were compiled by Giese (9) and by Rabideau (19). References listed therein trace the historical developments in the construction and use of visual acuity targets during the past half-century.

Some investigations of visual acuity published subsequently to the appearance of the former bibliography (9) have made use of factor analysis. The identification of several underlying factors in such studies suggests that, for possible future research, several acuity measuring devices should be so designed as to separately measure those factors, thereby allowing evaluation of the relative importance of such factors in activities requiring visual sensitivity.

Application of a Psychophysical Method

General conclusions stemming from other research in visual acuity have shown the superiority of certain target symbols and revealed the effect of certain factors, such as distance, illumination, brightness contrast, and exposure time (19).

In the present experiment these conditions were held constant, and particular attention was given to the selection of an extremely reliable measurement method, in order that differences between visual acuity thresholds on the various targets could be evaluated. Because it was found in an earlier investigation (23) that the Urban constant psychophysical method could be applied to the measurement of such thresholds, using as few as 50 trials per threshold and yielding reliability coefficients above .93, that measurement technique was selected.

[1] This monograph is a condensation of a dissertation at Purdue University (19). The author wishes to thank Professor Joseph Tiffin, chairman of his advisory committee, for his interest, guidance and encouragement; the other members of that committee, Professors E. J. Asher, N. C. Kephart, C. H. Lawshe, and W. V. Owen, for their helpful counsel; and his wife, Betty June, who assisted with the mechanics of writing this monograph.

Several underlying assumptions were necessary (17) to the application of a psychophysical method to this problem in visual acuity measurement. Briefly summarized, they are: (*a*) A threshold is an average value of an ever-changing sensitivity. (*b*) Both the judgments resulting in a threshold and the thresholds thereof must be considered as random samples of their respective infinite universes. (*c*) Sampling errors enter into either of the above sampling processes, as well as in the sampling of the subject (*S*) population.

Additional background on applied psychophysical methods in general may be found in Irwin (13), Jones (14), Preston (18), and Rich (20). Guilford (11) briefly treated the history of the constant methods and described the Müller-Urban method.

There have been several applications of the method of constant stimuli to technological problems similar in nature to the comparison of visual acuity thresholds obtained with different targets. One such paper (26) reports on the effect of floor tile size on apparent brightness of patterned tile, and the effect of cover markings on the apparent size of golf balls. Another (14) paper described the utilization of constant methods in photography and the determination of relationships between physical film emulsion and chemical variables and picture quality. In another article, Ferguson (6) described the statistical adaptation of the constant process to item selection, outlining a method of item selection based on the method of constant stimuli.

The Problem

The problem, restated more fully, was as follows: When visual acuity is reliably and precisely measured, using a single psychophysical measuring technique—the method of constant stimuli—and holding other conditions, such as lighting and distance, constant, and when thresholds are obtained on several different types of commonly used visual acuity targets:

1. What does an analysis of variance reveal about mean performances of subjects on the various targets?

2. How reliable are the thresholds obtained from those different targets, when a split-half method of reliability estimation is used?

3. How many factors underlie performances on the several targets' tasks?

Apparatus and Procedure

Target Design and Construction

The target series that was used consisted of eight different task groups, hereafter referred to as targets. Five of those targets consisted of Snellen letters. However, since one consisted of the Snellen "E," which was presented to *S* as a positional discrimination task, four of the targets were positional by nature of the task, and four were letter naming, or legibility targets.

The four positional targets shown in Fig. 1a, 1b, 1c, and 1d are the Ortho-Rater checkerboard design (also referred to as "C"), the Keystone Telebinocular dot-type test object ("K"), a Landolt circle ("L"), and the Snellen "E." In Fig. 1e, 1f, and 1g are found the letter target groups S_I, S_{II}, and S_{III}, consisting of the letters L, R, T, U, Z (S_I); C, F, P, S, V (S_{II}); and B, G, H, N, and Y (S_{III}). Target group S_{IV} consisted of a combination of all 15 of the above letters, which include 15 of the 17 different letters constituting the Snellen chart. (A Snellen chart published by Bausch and Lomb was used as a model for the construction of the experimental letter targets.)

The assignment of individual letters to target groups S_I, S_{II}, and S_{III} was made on the basis of a prior investigation (19) of letter legibility. Target group S_I consisted of five of the most legible letters; S_{II}, five of intermediate difficulty; and S_{III}, the five least legible letters.

The relative dimensions of the targets for any given target size were of two kinds:

1. Dimension of critical element. For the checkerboard design (C) the critical element was the width of the large-grid squares; for the dot design (K) it was the diameter of the dot; for the Landolt ring (L), the width of the break, or gap, and the width of the stroke; and for all Snellen letters (including "E"), the stroke width.

2. Over-all dimensions. The over-all dimensions of the large grid checkerboard target configuration were five times the critical element size, whereas the over-all dimensions of the checkerboard configuration, with all five grid patterns

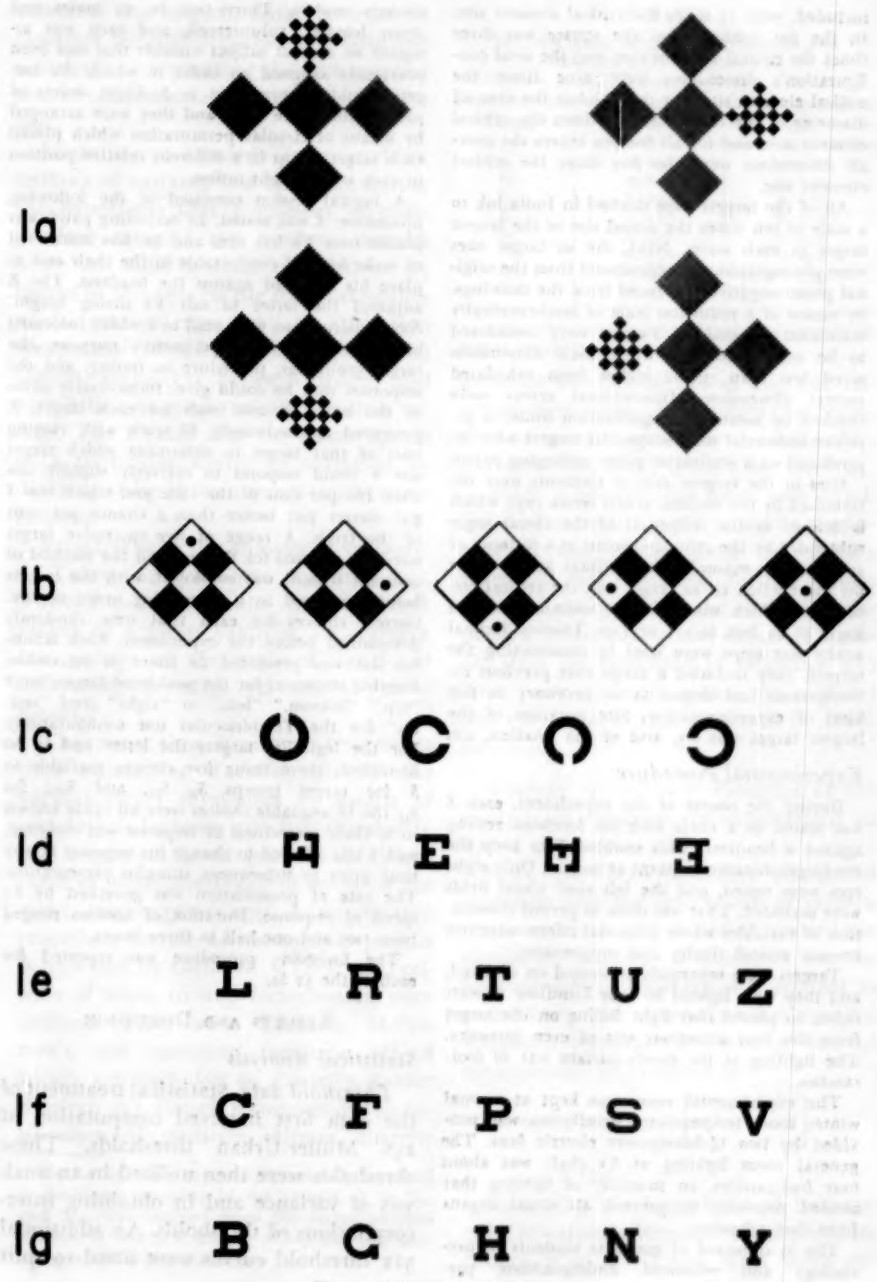

FIG. 1. Positional and letter targets.

included, were 15 times the critical element size. In the dot configuration the square was three times the critical element size, and the total configuration's dimensions were nine times the critical element size. For the Landolt the over-all diameter of the circle was five times the critical element size; and for all Snellen letters the over-all dimensions were also five times the critical element size.

All of the targets were drafted in India ink to a scale of ten times the actual size of the largest target in each series. Next, the 20 target sizes were photographically reproduced from the original photo negatives produced from the drawings, by means of a reduction scale of mathematically calculated dimensions. Targets were considered to be acceptable only when their dimensions erred less than $\pm.002$ inches from calculated correct dimensions. Dimensional errors were checked by means of magnification under a 40-power industrial microscope. All targets were reproduced on a semimatte photo enlarging paper.

Sizes of the targets' critical elements were determined by the decimal acuity series (23), which is defined as the reciprocal of the visual angle subtended by the critical element at a distance of 20 feet. For example, the decimal acuity value for the Snellen 20/20 target size, the critical element of which subtends one minute of visual angle at 20 feet, is 1/1 or 1.00. Twenty decimal acuity size steps were used in constructing the targets; they included a range that previous investigations had shown to be necessary to this kind of experimentation. Size notation of the largest target was 0.3, and of the smallest, 2.2.

Experimental Procedure

During the course of the experiment, each *S* was seated in a chair with his forehead resting against a headrest. This enabled *E* to keep the eye-target distance constant at 20 feet. Only right eyes were tested, and the left eyes' visual fields were occluded. That was done to permit elimination of variables whose potential effects were not known: retinal rivalry and suppression.

Targets were separately presented on an easel, and they were lighted by four Lumiline 40-watt tubes, so placed that light falling on the target from the four directions was of even intensity. The lighting at the easel's surface was 40 foot-candles.

The experimental room was kept at normal winter room temperature. Ventilation was provided by two ⅙-horsepower electric fans. The general room lighting at *S*'s chair was about four foot-candles, an intensity of lighting that seemed necessary to prevent *S*'s visual organs from dark adapting.

The *S*s consisted of graduate students in psychology and advanced undergraduate psychology majors. Thirty-two *S*s, 25 males and seven females, volunteered, and each was assigned an ordinal subject number that had been previously assigned an order in which the targets would be presented to *S*. Eight orders of presentation were used, and they were arranged by means of circular permutation which placed each target group in a different relative position in each of the eight orders.

A typical session consisted of the following procedure: *S* was seated, an occluding patch was placed over *S*'s left eye, and he was instructed to make himself comfortable in the chair and to place his forehead against the headrest. The *E* adjusted the latter to suit *S*'s sitting height. Next, instructions were read to *S* which informed him briefly of the experiment's purpose, the target groupings, procedure in testing, and the responses that he could give. Immediately prior to the series of test trials for each target, *E* presented approximately 40 trials with varying sizes of that target to determine which target size *S* could respond to correctly slightly less than 100 per cent of the time and which size *S* got correct just better than a chance per cent of the trials. A range of five successive target sizes was selected for the test and the method of constant stimuli was employed, with the targets being presented in a descending order of size. Correct choices for each trial were randomly determined before the experiment. Each stimulus size was presented 20 times in succession. Possible responses for the positional targets were "top," "bottom," "left," or "right" (and "center" for the Telebinocular dot configuration). For the legibility targets the letter had to be identified, there being five choices available to *S* for target groups S_{I}, S_{II}, and S_{III}; for S_{IV} the 15 available choices were all made known to *S*. Only correctness of response was recorded, and *S* was allowed to change his response at any time prior to subsequent stimulus presentation. The rate of presentation was governed by *S*'s speed of response. Duration of sessions ranged from two and one-half to three hours.

The foregoing procedure was repeated for each of the 32 *S*s.

Results and Discussion

Statistical Analysis

Threshold data. Statistical treatment of the data first involved computation of 256 Müller-Urban thresholds. These thresholds were then utilized in an analysis of variance and in obtaining intercorrelations of thresholds. An additional 512 threshold curves were fitted to split

halves of the original data, in order to compute reliability coefficients. Splitting was done by separating the first ten and second ten presentations of each stimulus size.

According to Edwards (5, p. 174), an analysis of variance is a means of testing the hypothesis that a number of independent samples are random samples from the same normal population. By means of this process the total variability of *Ss*' performances may be split into component parts, i.e., variability between targets, variability between orders, and variability between subjects within orders. The *F*-ratio tests the mean square of a particular source of variation by dividing it by its own mean-square-error estimate.

Subsequent to the analysis of variance, a Tukey analysis was applied (24). Tukey's process enables one to break means of a given variable, i.e., targets, into groups on the basis of significant differences between certain of those means.

The remaining statistical treatments were estimation of reliabilities of the various target thresholds by means of correlations between split-half thresholds, and a centroid factor analysis which was made of intercorrelations of seven[2] of the eight target groupings. A step-by-step procedure for the factor analysis process is provided by Guilford (11). Several criteria of when to stop factorization were utilized. They were Tucker's, McNemar's, and Saunders' formulas, all of which are included in Cattell's *Factor Analysis* (1). Cattell recommends the application of multiple criteria, thereby avoiding the fallibilities of any one criterion used alone. He believes that Saunders' criterion is probably the most effective criterion, and Tucker's the briefest.

[2] S_{IV}, a legibility test, which actually involved a combined grouping of the targets of S_I, S_{II}, and S_{III}, was omitted from the factor analysis because it was felt that it would not add anything to the factor matrices.

TABLE 1

MEAN VISUAL ACUITY THRESHOLDS AND STANDARD DEVIATIONS FOR ALL TARGETS (32 *Ss*)

Target*	Mean	*SD*
C	1.236	0.381
K	0.841	0.157
L	1.294	0.277
E	1.456	0.319
S_I	1.506	0.293
S_{II}	1.358	0.277
S_{III}	1.414	0.295
S_{IV}	1.254	0.230

* See the text for identification of the above target symbols.

Table 1 presents the mean and standard deviation of the visual acuity thresholds of the 32 *Ss*. Each individual's threshold was measured by the Urban constant process, with the thresholds expressed in decimal acuity units. (Threshold data are tabulated in reference 19.) The computational procedure utilized was the same as that outlined by Guilford (11).

Analysis of variance. Findings of the analysis of variance are summarized in Table 2.

Since the between-orders source of variation and the interaction of targets-versus-orders source failed to contribute significant *F* ratios, a Bartlett's test for homogeneity of variance was applied to each of the two error-variation sources. Bartlett's formula (5) for a χ^2 test of homogeneity of variance is:

$$\chi^2 = (2.3026)(n-1)\left[r(\log \frac{\Sigma s^2}{r})\right] - \Sigma \log s^2.$$

In testing the hypothesis of homogeneity of variance between-subjects-within-orders, a χ^2 of 11.340, with seven de-

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE FINDINGS

Source of Variation	Statistic			
	Sum of Squares (SS)	*df*	Mean Square (MS)	*F*
Between orders (*O*)	1.162	7	0.166	$\frac{MS_O}{MS_s} = 0.285$
Between *Ss* within orders	13.994	24	0.583	
Between targets (*T*)	9.582	7	1.369	$\frac{MS_T}{MS_{T\times s}} = 52.654^*$
Interaction $T \times O$	0.618	49	0.013	$\frac{MS_{T\times O}}{MS_{T\times s}} = 0.500$
Interaction $T \times Ss$ within orders	4.409	168	0.026	
Total	29.765	255	—	—

* Significant beyond the 1 per cent confidence level.

grees of freedom, was obtained. In the test for homogeneity of variance of the targets-versus-subjects-within-orders interaction, a χ^2 of 11.960 resulted, with the same number of degrees of freedom. In both cases the hypotheses of homogeneity of variance were considered to be tenable.

Tukey analysis of differences between means. Presented in Table 3 are mean thresholds for all targets, arranged in order of magnitude. Those means were subjected to a Tukey analysis (24). Included are differences between successive means. The differences were first subjected to a significance test against a *t*-ratio criterion set up for that difference between target threshold means which would be significant at the 1 per cent confidence level. Only the K and C target means differed significantly at the 1 per cent level. Targets S_I, E, S_{II}, S_{III}, L, S_{IV} and C, therefore, were called group *I*, and target K was designated group *II*.

A second step in the Tukey analysis was an application of straggler tests to the mean thresholds of the seven targets in group *I*, with the result that group *I* was split into two subgroups, *A*, and *B*, which differed significantly at the 1 per cent level.

Table 4 shows, under the heading "groups," the two groups of targets having threshold means which were signifi-

TABLE 3

TARGETS' THRESHOLD MEANS ARRANGED IN ORDER OF MAGNITUDE FOR TUKEY ANALYSIS, WITH SMALLEST AT LEFT. DIFFERENCES BETWEEN SUCCESSIVE MEANS

Statistic	Target							
	K	C	S_{IV}	L	S_{II}	S_{III}	E	S_I
Mean threshold	0.841	1.236	1.254	1.294	1.358	1.414	1.456	1.506
Difference	0.395	0.018	0.040	0.064	0.056	0.042	0.050	

TABLE 4

SUMMARY OF TUKEY ANALYSIS: SPLITTING OF TARGETS INTO SIGNIFICANTLY DIFFERENT GROUPS AND SUBGROUPS

Group*	Subgroup	Targets	Relative Difficulty
I	*A*	S_I, E, S_{III}	Easiest targets
	B	S_{II}, L, S_{IV}, C	Intermediate
II		K	Most difficult

* Probability that these groups and subgroups are samples from same normal population is less than .01.

cantly different. Table 4 also shows the significantly different subgroups of group *I*. Under "targets" are presented, in Table 4, the separate targets composing groups *IA*, *IB* and *II*.

Finally, an *F* test was applied to the between-means variability for the components in each subgroup. It was found that the between-means variation failed to yield an *F* significant at the 1 per cent level. Separation of targets into significantly different subgroups was terminated at that point. Three separable groupings and subgroupings were found, with the dot test being the most difficult, and certain of the Snellen letters the easiest-to-see targets.

Intercorrelations and factor analysis. Table 5 presents the intercorrelations of the target groups.

Following Guilford's (11) recommendation that the communalities of the correlational matrix should be estimated by successive approximations, two approximations were made of the factor loadings and communalities of the seven target groups. Table 6 shows the factor loadings and communalities obtained from each approximation.

Figure 2 graphically reveals the factor loadings of the seven tests. Although use of a greater number of test targets might have allowed a rotation of the axes, the lack of additional tests clustering around C (the checkerboard target) or K (the dot target) does not warrant a rotation of the axes which are shown in Figure 2. Naming of the factors is discussed in a subsequent paragraph.

Split-half reliabilities. Estimates of the reliabilities of thresholds obtained from the various targets are contained in Table 7. The top row represents the raw split-half coefficients, and the bottom represents those coefficients yielded by the Spearman-Brown prophecy formula for a test having the length of the one used. All the Spearman-Brown coefficients were found to be .90 or higher.

The thresholds that were computed for

TABLE 5

INTERCORRELATIONS OF THRESHOLDS YIELDED BY EIGHT TARGETS

Target	K	L	E	S_I	S_{II}	S_{III}	S_{IV}
C	.381	.584	.616	.511	.540	.626	.594
K		.637	.718	.776	.756	.714	.693
L			.873	.813	.851	.871	.837
E				.871	.900	.921	.907
S_I					.909	.870	.880
S_{II}						.935	.852
S_{III}							.859

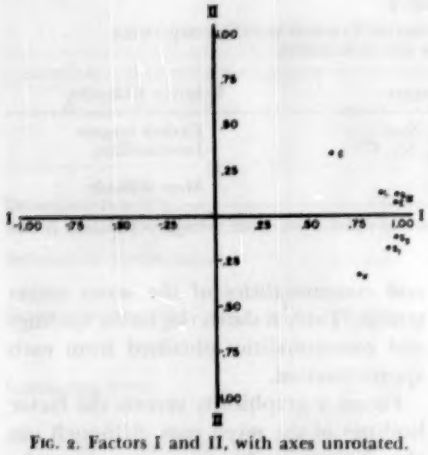

FIG. 2. Factors I and II, with axes unrotated.

purposes of split-half reliability estimation are tabulated in reference 19.

Correlation of target area with the mean thresholds of the various targets. During the computational processes it was observed that there was a consistent tendency for certain of the targets to yield higher thresholds than did others. An additional hypothesis was formulated to the effect that fineness of threshold was positively related to the over-all amount, or area in square minutes of visual angle, of the black portions constituting the critical elements of standard-sized targets, when the standard size notation is 1.0 (decimal acuity series). It should be noted that these are relative areas, since they were computed with target size held constant.

In order to test the above hypothesis, the thresholds and areas were first ranked on the basis of total blacked-in area of each target's critical elements. Areas and ranks for the different targets are as follows: E, 16.0, 1; S_{III}, 14.0, 2; S_I, 13.7, 3; S_{IV}, 13.6, 4; S_{II}, 13.2, 5; C, 13.0, 6; L, 11.6, 7; and K, 0.8, 8.

Correlations were computed using a rank-difference procedure between all targets, and between positional targets only. Since the letter groups' areas were average values dependent upon the arbitrary assignment of letters to those groups, a separate correlation between the foregoing ranks was not computed for the legibility targets. Table 8 shows the area and the threshold rankings for each of the above correlations. Rho values obtained were .81 for all targets, and .80 for positional targets. The former is significantly greater than zero, according to Kendall (15), with confidence level at .008. Because of the small target n of the positional target coefficient, its significance could not be determined by means of available tables.

Discussion of Findings

That the between-targets source of variation yielded the only significant F ratio in the analysis indicates the tenability of several conclusions. First, the various ex-

TABLE 6

FIRST AND FINAL APPROXIMATION OF FACTOR LOADINGS AND COMMUNALITIES

Test	First Approximation			Final Approximation		
	h_1	h_2	h_i^2	h_1	h_2	h_i^2
C	.636	.369	.540	.624	.348	.510
K	.776	−.351	.725	.770	−.334	.705
L	.901	.134	.830	.896	.121	.818
E	.953	.069	.913	.954	.069	.915
S_I	.927	−.208	.902	.928	−.191	.897
S_{II}	.954	−.152	.933	.956	−.118	.928
S_{III}	.958	.123	.933	.961	.103	.935

TABLE 7
SPLIT-HALF RELIABILITY COEFFICIENTS

Coefficient	Target							
	C	K	L	E	S_I	S_{II}	S_{III}	S_{IV}
r_{11}	.896	.828	.956	.818	.963	.890	.845	.935
r_{11}*	.945	.906	.977	.900	.981	.943	.916	.967

* Coefficient r_{11} is the Spearman-Brown r.

perimental target objects cannot be considered to constitute samples of the same population. The Tukey analysis substantiated that, indicating that the experimental targets actually consist of subgroups of targets, differing significantly in difficulty.

Second, the order in which targets were presented did not affect performances significantly. Therefore, it may be concluded that practice and fatigue effects throughout the two-and-one-half hour testing sessions were either negligible, or they cancelled each other's effect.

Third, the target-order interaction failed of significance, indicating that *Ss*' performances on the various targets were not differentially affected by the order in which those targets were presented.

Since Bartlett's tests of the assumption of homogeneity of variance indicated the tenability of the homogeneity assumption, a conversion of the thresholds was not necessary.

The confidence level used throughout this investigation may be characterized as high. For example, had a 5 per cent level been used in the Tukey analysis, it is likely that almost all of the targets would have been found to differ significantly from each other. The error, if any, resulting from use of the 1 per cent confidence level has been one in the direction of reducing the number of independent target groupings. It is interesting to note that the separation of Snellen letters into more homogeneous groupings on the basis of legibility changed the relative difficulty of those groups. The "easiest" and "most difficult" groups were significantly less difficult than the "intermediate" group. A pertinent conclusion to this finding would probably be that those persons who would construct letter charts on the basis of the legibilities of letters should proceed with caution. Legibility will vary with the total number and varieties of letters from which the testee can select in making his response. Legibility of any letter is thus a function of

TABLE 8
MEAN THRESHOLD AND AREA RANKS FOR: (*a*) ALL TARGETS, AND (*b*) POSITIONAL TARGETS ONLY

Target Group	Variable Ranked*	Target							
		C	K	L	E	S_I	S_{II}	S_{III}	S_{IV}
All targets	Area	6	8	7	1	3	5	2	4
	Threshold	7	8	5	2	1	4	3	6
Positional targets	Area	2	4	3	1				
	Threshold	3	4	2	1				

* Lowest number represents the highest rank, which in turn is assigned to the largest area and to the smallest threshold.

the Gestalt of which it is a part.

The findings of the present investigation tend to substantiate Cowan (4), who concluded that a black dot on a white ground suffers from irradiation of the white, yielding higher threshold measurements than do letters. The evidence tends to lend support to Imus' (12) conclusion that Ortho-Rater tests are significantly more difficult than letter tests.

One may generally conclude that a test composed of Snellen letters differs in difficulty from a dot-type positional test or from a tumbling "E" test. Since those thresholds are not likely to be equivalent in individual cases, extreme caution should be urged in the substitution of the scores from one such test for the scores on another. To put it differently, a visual acuity rating of 20/20 on the tumbling "E," the Snellen letter chart, and a dot-type test means, in each case, a different level of acuity. Thus, the estimation of a patient's visual acuity is dependent upon the type of measuring device the clinician chooses to use.

Conclusions by Fink, Ferree, and Rand (7), and Giese (9), that Snellen letters should be discarded tend to be supported by this study's findings that threshold is a function of the letter combination.

Comparison of the factor analysis loadings with those found by the Army (27) and by Cook (3) indicates that a single major factor loading is present in all target performances. In both of the above investigations that major factor was named "retinal resolution." The second factor axis in the present analysis probably represents a brightness discrimination factor. If rotation were possible in Fig. 2, the probable result would be to reduce the checkerboard test (C) loading in factor II to zero, as was the case in the Army study. All other targets would show increases in factor II loadings. According to the Army study (27), the Ortho-Rater checkerboard target appears to be the purest test of retinal resolution, and the dot-type test, the purest measurer of brightness discrimination.

Additional research in development of visual test objects and factor analyses of test performances on them appears to be advisable. It is possible that visual acuity will one day be thought of as a group of performance traits. Validation of relatively pure measures of certain visual acuity factors awaits the construction of such measuring devices. Prediction of performance on different tasks requiring differing amounts of different visual acuity traits could be clinically as well as industrially useful.

With regard to threshold reliabilities, earlier findings (23) appear to be verified. A method of constant stimuli yields high reliabilities with as few as 100 stimulus presentations per test. Differences between the coefficients, which range from .90 to .98, may not be significant. The two lowest r's, however, were those obtained using the dot test and the tumbling "E," whereas the two highest were one of the letter target groups (S_4) and the Landolt ring.

Cowan's conclusion (4) that the Landolt is inferior in reliability to letter targets is not tenable in light of the findings of the present investigation. Fink's (8) conclusion that the tumbling "E" is inferior in reliability to the Landolt circle appears to be corroborated.

Evaluation of the finding that target area is related to threshold is difficult, since the hypothesis was evolved from the original experimental data. It is suggestive, however. If further research corroborates that finding, it will definitely point out the need for standardizing the relative dimensions of visual test objects. For instance, if a 3×3 black square

checkerboard should yield significantly different visual acuity thresholds from one having 5×5, or 25, black squares as critical elements, then the two must be considered to be measuring the same thing in different ways. The work of certain constructors of test objects, e.g., letters having other than the Snellen one- and five-minute dimensions (25), would be largely invalidated.

Summary and Conclusions

An experiment was conducted, using 32 *Ss*, to compare visual acuity measurements obtained with different types of commonly used visual acuity targets. The psychophysical method of constant stimuli was used. Variables such as distance and lighting were held constant. The threshold data were subjected to an analysis of variance, a factor analysis, and estimation of threshold reliabilities.

Results of the statistical analysis showed that visual acuity measurements differed significantly as a function of target, but not as a function of an interaction of practice and fatigue or of the order of presentation. At least two underlying factors determined performances on the various tests; these factors were named retinal resolution and brightness discrimination. Split-half reliability estimates indicated that the measurement method yields highly reliable measures for all targets (above .90). Relative area of black present in critical elements of targets of the same size was found to correlate positively with fineness of threshold.

Conclusions that appear to be supported by the findings are:

1. Visual acuity thresholds are dependent upon test object design and nature of the acuity measurement task.

2. The meaningfulness of Snellen letter thresholds is limited by the finding that letter difficulty, and hence threshold, is a function of what letters are combined in the test.

3. Snellen letter targets yield finer thresholds than do nonletter targets.

4. The number of letters involved as test choices is negatively related to fineness of threshold of Snellen letter tests.

5. Neither incremental conditions, such as practice, nor decremental ones, like fatigue, acted to significantly affect the measurements obtained on the various targets.

6. Caution should be exercised in comparing visual acuity measurements obtained with different types of targets.

7. Findings of a factor analysis indicate that there are two (or more) kinds of visual acuity.

8. Test objects designed to measure as purely as possible the basic abilities indicated by factor analysis would be of value in clinical and industrial testing.

9. It may be concluded that the method of constant stimuli and Urban constant process yields highly reliable visual acuity thresholds with as few as 100 stimulus presentations per test.

10. There is some evidence to indicate that thresholds of visual acuity are related to area of black composing the critical elements of the various targets used in this study.

Further research that is indicated by the present investigation falls into two areas:

1. The development of practical visual acuity test targets separately measuring the various factors. Such tests could be clinically and statistically validated, and they could prove of value in industrial personnel selection.

2. The relationship of area of target design to threshold should be investigated, with a view toward standardization of test object dimensions.

REFERENCES

1. CATTELL, R. B. *Factor analysis.* New York: Harper, 1952.
2. COOK, E. B. *A factor analysis of acuity and phoria measurements obtained with commercial screening devices and by standard clinical methods.* New London, Conn.: Naval Medical Research Laboratory, 1948. (Progr. Rep., No. 4, Project NM-003-011 [X-493].)
3. COOK, E. B. *Comparative performance of commercial screening devices and far and near wall charts utilizing the same test targets.* New London, Conn.: Naval Medical Research Laboratory, 1948. (Rep. No. 5, Project NM-003-011 [X-493].)
4. COWAN, A. Test letters which comply with the physiological requirements of a visual test object. *Amer. J. Ophthal.*, 1928, **11**, 625-628.
5. EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
6. FERGUSON, G. A. Item selection by the constant process. *Psychometrika*, 1942, **7**, 19-29.
7. FERREE, C. E., & RAND, G. A new method of rating visual acuity. *J. gen. Psychol.*, 1941, **25**, 143-176.
8. FINE, W. H. An evaluation of visual-acuity symbols. *Amer. J. Ophthal.*, 1945, **28**, 701-711.
9. GIESE, W. J. The interrelationship of visual acuity at different distances. Unpublished doctor's dissertation, Purdue Univer., 1945.
10. GIESE, W. J. The interrelationship of visual acuity at different distances. *J. appl. Psychol.*, 1946, **30**, 91-106.
11. GUILFORD, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.
12. IMUS, H. A. Comparison of Ortho-Rater with clinical ophthalmic examinations. *J. Aviat. Med.*, 1949, **20**, 2-23.
13. IRWIN, F. W. Psychophysical measurement methods. *Psychol. Bull.*, 1935, **32**, 140-171.
14. JONES, L. A. Psychophysics and photography. *J. opt. Soc. Amer.*, 1944, **34**, 66-68.
15. KENDALL, M. G. *Rank correlation methods.* London: Chas. Griffin & Co., Ltd., 1948.
16. NEWHALL, S. M. Linear interpolation vs. the constant process. *Amer. J. Psychol.*, 1927, **38**, 390-402.
17. POSTMAN, L., & BRUNER, J. S. The reliability of constant errors in psychophysical measurement. *J. Psychol.*, 1946, **21**, 293-299.
18. PRESTON, M. G. Psychophysical measurement methods. *Psychol. Bull.*, 1938, **35**, 63-83.
19. RABIDEAU, G. F. Differences in visual acuity measurements obtained with different types of targets. Unpublished doctor's dissertation; Purdue Univer., 1953.
20. RICH, G. J. Psychophysical measurement methods. *Psychol. Bull.*, 1925, **22**, 613-648.
21. SLOAN, Louise L. Measurement of visual acuity: a critical review. *A.M.A. Arch Ophthal., Chicago*, 1951, **45**, 704-725.
22. SNELLEN, H. *Probebuchstaben zur Bestimmung der Sehscharfe.* Utrecht: P. W. van de Weijer, 1862.
23. TIFFIN, J., & RABIDEAU, G. F. Reduced trials and reliability of visual acuity thresholds obtained by the method of constant stimuli. *J. appl. Psychol.*, 1955, **39**, 78-81.
24. TUKEY, J. W. Comparing individual means in the analysis of variance. *Biometrics*, 1949, **5**, 99-114.
25. WAUGH, D. D. A combination of the Snellen and the Landolt test types. *A.M.A. Arch. Ophthal., Chicago*, 1939, **21**, 671-672.
26. WENDT, G. R. Two industrial applications of a psychophysical method. *J. appl. Psychol.*, 1932, **16**, 269-276.
27. *Studies in visual acuity; PRS Report No. 742.* Department of the Army, The Adjutant General's Office. Washington, D.C.: U.S. Government Printing Office, 1948.

(Accepted for publication October 21, 1954)